

A INTERPRETAÇÃO DE MEDIDAS DE TENDÊNCIA CENTRAL DE FUTUROS PROFESSORES E EDUCADORES NA REALIZAÇÃO DE UMA INVESTIGAÇÃO ESTATÍSTICA

Raquel Santos

Unidade de Investigação do Instituto de Educação da Universidade de Lisboa
raquelfms@gmail.com

João Pedro da Ponte

Instituto de Educação da Universidade de Lisboa
jpponte@ie.ul.pt

Resumo

Com a inclusão do ensino da Estatística nos primeiros anos da educação básica torna-se necessário fazer uma aposta ainda mais forte na formação inicial de professores e educadores neste tema. Consequentemente, é importante compreender os conhecimentos que os futuros professores e educadores de infância possuem tanto em Estatística como na sua didática. Com esse objetivo, analisamos os relatórios que elaboraram no âmbito de uma investigação estatística, procurando compreender que significados os futuros professores e educadores atribuem às medidas de tendência central. Referimos, ainda, implicações dos resultados do estudo para a formação inicial de professores.

Palavras-chave: Estatística, Média, Mediana, Moda, Investigação estatística, Formação inicial.

Introdução

Nas *Orientações Curriculares para a Educação Pré-Escolar* (ME, 1997) e no *Programa de Matemática do Ensino Básico* (ME, 2007), é dada ênfase à Estatística desde os primeiros anos. No ensino deste tema assume um lugar central o desenvolvimento da literacia estatística, isto é, a capacidade de interpretar, avaliar criticamente e comunicar acerca de informação e mensagens estatísticas (Gal, 2002), vertente fundamental na educação de todo o cidadão e, ainda mais, na educação do futuro professor. Neste quadro, será importante saber de que modo os futuros professores e educadores de infância interpretam e comunicam os resultados estatísticos. Esta comunicação relata parte de uma investigação que procura compreender o conhecimento de Estatística e da sua Didática de futuros professores e educadores de infância quando terminam a sua licenciatura em Educação Básica. Debruçamo-nos sobre os relatórios de investigações estatísticas, analisando o modo

como os futuros professores e educadores interpretam as três medidas de tendência central.

Enquadramento teórico

A realização de investigações estatísticas permite aos alunos tornarem-se ativos na aprendizagem da Estatística. Nestes projetos, os alunos escolhem um tema do seu interesse, definem objetivos, escolhem instrumentos de recolha, selecionam amostras, recolhem, codificam, analisam e interpretam dados para responder às questões propostas (Batanero & Godino, 2005). Durante uma investigação percorrem todos os passos do ciclo PPDAC (Problema, Plano, Dados, Análise e Conclusões) de Wild e Pfannkuch (1999), num ambiente propiciador de uma aprendizagem significativa (Ponte, 2007), podendo apreciar a importância e a dificuldade do trabalho estatístico e o interesse da Estatística na abordagem de problemas da vida real (Batanero & Godino, 2005). Além disso, o trabalho com investigações estatísticas permite conhecer os pontos fracos do conhecimento matemático dos alunos e, por vezes, até conceitos e ideias que pareciam bem consolidados (Ponte, 2007).

O professor é um elemento central do ensino, tendo de saber muito bem a matéria para a poder ensinar (Ma, 1999). Além disso, o professor tem também de possuir um conhecimento didático da Matemática (o *pedagogical content knowledge* de Shulman, 1986). Para Mulekar (2007), os futuros professores necessitam de uma compreensão aprofundada de conceitos estatísticos de modo a darem um significado coerente aos resultados. As medidas de tendência central (moda, média e mediana) têm uma relevância particular por se encontrarem frequentemente no dia-a-dia (Groth, 2006). Para Groth (2006), a compreensão destas medidas é uma componente importante da literacia estatística. No entanto, segundo Jacobbe (2008), mesmo professores do 1.º ciclo exemplares não possuem conhecimento conceptual de dois conceitos básicos da Estatística – moda e média.

Em Portugal, a moda aparece no novo programa de Matemática nos 3.º e 4.º anos (ME, 2007) e os professores tendem a considerar que se trata de um conceito de fácil compreensão. No entanto, alguns estudos sugerem o contrário. Assim, Fernandes (2009), num levantamento de dificuldades e erros em Estatística de alunos dos 7.º e 12.º anos e de futuros professores dos 1.º e 2.º ciclos, refere falhas na compreensão deste conceito, quando selecionam a maior frequência em vez do valor da variável

correspondente. Num estudo com 40 futuros professores na licenciatura em Educação Básica foram frequentes as respostas do tipo “a moda é 9, pois existe maior número de alunos que vê televisão” (Martins, Pires & Barros, 2009, p. 7). Na interpretação deste conceito é frequente associar a moda “ao maior número que incide numa tabela de resultados”, à maior frequência absoluta, ao “valor que aparece mais vezes” e à “categoria ou classe de maior frequência” (Martins, Pires & Barros, 2009, p. 11). Algumas destas interpretações demonstram um conhecimento confuso, dado que o valor que aparece mais vezes pode ser visto como o número na coluna das frequências absolutas que se repete e não o valor da variável que se repete mais vezes.

O conceito de média é introduzido no 2.º ciclo (ME, 2007). A investigação sobre este conceito é bastante vasta, dado ser muito utilizado em estudos estatísticos. Leavy e Loughlin (2006) indicam que existem dois tipos de compreensão – conceptual e processual. Processualmente, a média aritmética é o valor no qual a soma dos desvios em relação à média numa direção é igual à soma dos desvios na outra direção. Conceptualmente, a média pode ser vista como ponto de equilíbrio ou centro de gravidade, representando o conjunto de dados. Para os autores, interpretações da média como distribuição equitativa dos dados (ou seja, o valor que representa o conjunto de dados como se todos eles fossem iguais), ou como valor de equilíbrio, onde os valores mais altos compensam os valores mais baixos, demonstram compreensão conceptual do conceito. No cálculo da média, um erro frequente é determinar a média das frequências absolutas de variáveis qualitativas (Martins, Pires & Barros, 2009). Os alunos evidenciam diversas interpretações do conceito de média, nomeadamente a atribuição de significado com base no algoritmo de cálculo (Chatzivasileiou, Michalis & Tsaliki, 2010; Fernandes & Barros, 2005), com respostas do tipo “soma de todos os resultados divididos pelos valores existentes”, ou “soma dos números” (Martins, Pires & Barros, 2009). Outros associam a média à “noção de equilíbrio”, ao “valor médio”, ao “valor que equilibra os valores mais altos e mais baixos” (Martins, Pires & Barros, 2009), ao valor baseado na distribuição equitativa dos dados (“*fair share*”), ao valor que se deve ter se todos os dados fossem iguais, ao valor típico esperado ou à medida de localização (valor próximo mas não exato) (Chatzivasileiou, Michalis & Tsaliki, 2010). Como interpretações incorretas temos respostas baseadas no valor máximo, no valor mínimo, num dado específico, na mediana e na moda (Chatzivasileiou, Michalis & Tsaliki, 2010; Leavy & Loughlin, 2006). Nestes últimos casos, os alunos falham no reconhecimento

do conjunto de dados como um todo e tendem a focar-se em valores individuais (Chatzivasilieiou, Michalis & Tsaliki, 2010).

Relativamente à mediana, conceito do 3.º ciclo (ME, 2007), o panorama também é problemático. Foram detetadas dificuldades na compreensão e interpretação deste conceito tanto em alunos do 12.º ano (Fernandes, 2009), como em futuros professores dos 1.º e 2.º ciclos (Fernandes & Barros, 2005). No seu cálculo, muitos futuros professores indicam o valor central das frequências absolutas, outros confundem este conceito com a moda, e o erro mais frequente é determinar o valor central sem ordenar os dados (Martins, Pires & Barros, 2009). Mesmo professores do 1.º ciclo com muita experiência calculam a média quando lhes é pedida a mediana (Jacobbe, 2008). Na interpretação deste conceito, há associações ao valor central (mas nem sempre da forma mais correta) bem como ao “valor que está no meio”, ao valor que “divide ao meio”, ao valor “intermédio”, ao valor “que está mais ou menos no meio”, ao “quartil 2” sem mais explicações, ao valor que “divide a amostra ao meio e equilibra os grandes valores com os pequenos”, ao “número médio de todos os resultados” e ao “valor médio” (Martins, Pires & Barros, 2009). As três últimas interpretações mostram alguma confusão entre os conceitos de mediana e média.

Metodologia

Os participantes deste estudo são futuros professores e educadores de infância da Licenciatura em Educação Básica de uma Escola Superior de Educação que frequentaram a unidade curricular Tópicos de Matemática Discreta, Estatística e Probabilidades no 2.º semestre do 2.º ano, em 2010/11. Nesta unidade curricular, a única da licenciatura dedicada ao desenvolvimento do conhecimento estatístico, trabalharam os conceitos estatísticos numa ótica de descoberta, tendo sido dada grande ênfase à sua interpretação em contexto real através de exercícios e problemas. Foi-lhes também proposta a realização de uma investigação estatística, em grupo, de um tema por eles escolhido, mas com possível utilização com futuros alunos. Durante a realização da investigação, foi-lhes pedido o registo de todo o processo num relatório redigido também em grupo e apresentado à turma no final do semestre, incluindo uma análise e interpretação de dados, uma conclusão sobre o tema e uma reflexão sobre o trabalho realizado. Esta comunicação analisa estes relatórios escritos de 36 futuros professores que frequentaram a unidade curricular no regime de avaliação contínua (18

em regime diurno e 18 em regime pós-laboral), tendo em vista perceber o seu conhecimento sobre medidas de tendência central. Uma vez que os participantes recorrem ao programa Excel para calcular as medidas, a ênfase é colocada na interpretação dessas medidas no contexto de cada questão. Nesta parte do estudo foi adotada uma metodologia quantitativa com a elaboração de grupos com as diferentes interpretações observadas nos relatórios. Por limitações de espaço, discutimos apenas resultados que consideramos mais pertinentes, com destaque para resultados que não foram observados em estudos anteriores.

Os futuros professores organizaram-se em dezasseis grupos e escolheram temas variados, como por exemplo reciclagem, desporto, alimentação, consumo de água e rotinas diárias elaborando questionários com um mínimo de 11 questões e um máximo de 25. No total foram feitas 273 questões, das quais 81 relativas a variáveis quantitativas e as restantes 192 referentes a variáveis qualitativas. No tratamento dos dados todos os grupos utilizaram o programa Excel, recorrendo às respetivas funções para calcular a média e a mediana. Para o cálculo da moda perceberam ser mais conveniente não recorrer à função “moda” do Excel, uma vez que esta não funciona para variáveis qualitativas nem para o caso de uma variável ter mais do que uma moda.

Resultados

Moda

Surgem muitas interpretações do conceito de moda nos relatórios das investigações estatísticas (Tabela 1). Assim, cinco dos grupos associam a moda a algo que acontece “mais vezes”, como no exemplo da análise da questão “Qual a sua profissão?” onde o grupo refere “a moda do grupo de profissões é “profissões intelectuais e científicas”, pois é o grupo de profissões que se repete mais vezes na amostra” (RG4D, p. 11). Também associadas a uma resposta que aparece muitas vezes, mas desta vez aludindo à “frequência”, ou palavras da família, foram as interpretações feitas por onze grupos. O exemplo “a moda (...) é casa, pois é o valor do gráfico e da tabela mais elevado” na análise da questão “Onde toma o pequeno-almoço?” (RG1PL, p. 29) só não criou confusão por se tratar de uma variável qualitativa.

Três dos grupos utilizaram interpretações que se referiam a algo que predomina como na questão “Quantas pessoas tomam banho diariamente em sua casa?” que suscitou a seguinte análise “é possível constatar que o número predominante é de 3 banhos com

cerca de 42%” (RG5D, p. 14). Praticamente todos os grupos usaram expressões como “a maior parte” ou expressões semelhantes que traduzem uma ideia correta: “a maior parte dos meninos vai para o ATL” depois da escola (RG6D, p. 23).

Interpretações da moda como a seguinte, feitas por oito grupos, são muito ligadas ao contexto, fazendo uso da expressão “mais” ou “maior”: “o transporte mais utilizado (...) é o carro” (RG6PL, p. 9). Um grupo mencionou ainda a moda como sendo algo comum na amostra: “o jogo mais comum entre estes alunos é o futebol” (RG2PL, p. 15). Esta expressão mostra que o grupo olha para a moda como sendo o valor típico expectável, interpretação muitas vezes associada à média (Chatzivasileiou, Michalis & Tsaliki, 2010).

Tabela 1 – Interpretações adequadas da moda

Interpretações adequadas	# grupos
Valor/modalidade referido(a)/repetido(a)/verificado(a)/escolhido(a)/que surge mais vezes	5
Valor/modalidade associado(a) à maior frequência	11
Grande parte da amostra...	2
Maior parte da amostra.../a (grande) maioria da amostra.../...maioritariamente...	14
Valor/modalidade associado(a) a mais ou a um maior número de inquiridos	5
Outra interpretação usando “mais” ou “maior” em contexto	8
Valor/modalidade predominante/valor/modalidade que mais predomina/...predominantemente...	3
Valor/modalidade com mais relevância	1
Valor/modalidade mais comum	1
Valor/modalidade satisfatório(a)	1
Valor/modalidade preferido(a)	2

Entre os problemas de interpretação (Tabela 2), na análise da questão “Desde quando tem o hábito de reciclar?” (RG2D, p. 9), um grupo usou a expressão “notámos que 36% das pessoas começaram a reciclar entre 2006 e 2008, em que, sendo esta a classe modal, concluímos que grande parte da amostra começou a reciclar nesta altura”. Esta resposta é problemática porque o facto de acontecer algo numa grande parte da amostra, não significa que essa “parte” seja a maior e, por conseguinte, referente à moda.

Adicionalmente, três grupos fizeram afirmações gerais o que também é problemático, uma vez que a moda não é o que acontece a todos os inquiridos, mas a uma maioria. É ainda de referir um grupo que fez a seguinte afirmação: “as idades a que realizamos mais inquéritos foi a jovens com 19 anos” (RG1D, p. 8). Nota-se alguma confusão, não sabemos se a nível de compreensão da moda, se a nível de construção frásica e português. Esta frase transmite a ideia de que se realizaram inquéritos a idades e não a pessoas, o que pode ser confuso para quem lê os relatórios.

Tabela 2 – Problemas de interpretação da moda

Problemas de interpretação	# grupos
Grande parte da amostra...	1
Generalização	3
Confusão	1
Sem interpretação (variáveis quantitativas)	13
Sem interpretação (variáveis qualitativas)	15

De destacar que, no caso das variáveis qualitativas, quinze grupos não fizeram qualquer referência, nem mesmo implícita, ao conceito de moda e ao que esta medida representa e significa. O mesmo aconteceu com treze grupos para as variáveis quantitativas. Isto mostra que, principalmente em variáveis onde a moda é a única medida estatística que se pode determinar, os participantes limitam-se a uma leitura dos gráficos e tabelas.

Média

Em termos de interpretações adequadas do conceito de média (Tabela 3), um grupo referiu a seguinte expressão: “a média das idades é de 3,4, que nos representa o equilíbrio das idades” (RG1PL, p. 11). A resposta demonstra que o grupo percebe que a média pode ser vista como um ponto de equilíbrio, mas não deixa transparecer de que modo compreende realmente o significado desta medida. Apesar de esta interpretação poder demonstrar uma compreensão conceptual, não é evidente o que os participantes compreendem deste conceito. Houve ainda um grupo que considerou a média como o valor “normal”, ou seja, o que é normal acontecer. Esta interpretação vai ao encontro da interpretação de valor típico esperado, considerada correta mas, mais uma vez, não demonstra compreensão do conceito.

Tabela 3 – Interpretações adequadas da média

Interpretações adequadas	Número total de grupos
Valor que representa o equilíbrio	1
Valor “normal”	1
Distribuição equitativa	2

Também no caso da média foram observadas interpretações pouco adequadas que sobretudo revelam falhas na distinção entre a média e outra medida estatística (Tabela 4). Um grupo fez confusão entre a média e as medidas de dispersão, usando a média para indicar se os dados estão ou não dispersos: “os valores da média é 7,6 (...), logo podemos concluir que os valores não são muito dispersos” (RG6D, p. 28). De realçar que nenhum dos grupos, em algumas das variáveis quantitativas, interpretou a média. Nestes casos estão incluídas as afirmações do tipo “a média é...” ou “a amostra, em média...” por considerarmos que estas frases não acrescentam qualquer sentido ao valor da média.

Tabela 4 – Problemas de interpretações da média

Problemas de interpretação	# grupos
Confusão entre média e outras medidas de tendência central	1
Confusão entre média e medidas de dispersão	1
Sem interpretação	16

Mediana

No que respeita a interpretações corretas do conceito de mediana (Tabela 5), há a registar a seguinte interpretação de um dos grupos: “até à mediana estão 50% da amostra e para lá da mediana também estão” (RG5D, p. 13), o que demonstra que estes participantes compreendem que a mediana divide a amostra ao meio, 50% para cada lado da mediana, podendo no entanto ser problemático a não inclusão do valor da mediana na segunda metade da amostra (“para lá da mediana”).

Tabela 5 – Interpretações adequadas da mediana

Interpretações adequadas	# grupos
Valor que divide os dados ordenados	2
Até esse valor estão 50% da amostra	1
50% dos inquiridos... ou menos...	3
50% dos inquiridos... no máximo...	1

De realçar outras interpretações adequadas que revelam compreensão deste conceito no contexto dos dados analisados usadas por quatro grupos, do tipo: "50% das pessoas pretende regressar à feira, no mesmo ano, no máximo 2 vezes" (RG3D, p. 31). Este tipo de interpretação mostra uma compreensão profunda do que significa a mediana no contexto da variável em estudo, recorrendo a expressões que na realidade fazem sentido na interpretação dos dados e que vão para além do abstrato.

No entanto, nem tudo foram interpretações adequadas (Tabela 6). Quatro grupos atribuíram significados que revelam compreensão que a mediana é um valor central e que divide os dados em 50% para cada lado, mas que não mostram grande compreensão do conceito, evidenciando alguma confusão: "o quartil 2 é 50% dos dados" (RG4PL, p. 11). Nesta interpretação escrita o grupo confunde-se e refere que a mediana é o valor que tem 50% dos dados, o que não é correto.

Tabela 6 – Problemas de interpretação da mediana

Problemas de interpretação	# grupos
Valor que está no centro	1
Valor equivalente a 50%	1
Valor que é 50% dos dados	1
Valor para um máximo de 50% dos inquiridos	1
Confusão entre mediana e moda	1
Confusão entre mediana e medidas de dispersão	1
Sem interpretação	15

O seguinte exemplo mostra que o grupo tentou fazer uma formulação do tipo realçado nas interpretações adequadas, mas fez confusão, talvez por não compreender bem o conceito: "50% das pessoas no máximo vêm ao festival com 3 acompanhantes" (RG3D, p. 23). Outras interpretações inadequadas mostram uma certa confusão entre a mediana

e outras medidas estatísticas. De salientar o caso em que o grupo usa a mediana para concluir sobre a dispersão dos dados: “mediana é 8, logo podemos concluir que os valores não são muito dispersos” (RG6D, p. 28). É de realçar ainda que só um grupo fez a interpretação do conceito de mediana para todas as variáveis quantitativas que estudou.

Conclusões e implicações

A análise dos relatórios dos 16 grupos evidencia falhas na compreensão das medidas de tendência central dos futuros professores e educadores. Mesmo quando analisam estes conceitos no contexto dos dados recolhidos, atribuem significados às medidas que mostram alguma confusão relativamente ao seu significado, por vezes não fazendo distinção entre diferentes medidas estatísticas. Adicionalmente, talvez por terem de realizar um relatório escrito com a análise de diversas variáveis e por não quererem repetir sempre a mesma expressão, utilizaram interpretações não mencionadas por outros investigadores como referir a moda como algo que acontece “a grande parte”, “à maior parte” e “a um grande número” de inquiridos (interpretações adequadas) ou como valor relevante ou satisfatório e as generalizações (interpretações não adequadas). Alguns grupos atribuíram significados apropriados, como a média sendo o ponto de equilíbrio (Leavy & Loughlin, 2006), mas sem deixar transparecer o modo como compreendem realmente o significado das medidas. Relativamente à mediana observou-se a interpretação concreta e contextual que 50% da amostra está associada a este valor ou menos. Assim, as interpretações inseridas no contexto dos dados a tratar são as que demonstram maior compreensão dos conceitos. Para além disso, uma interpretação pode ser considerada adequada ou não, dependendo tanto do tipo de variável (quantitativa ou qualitativa) como da informação associada (por exemplo se refere a frequência no caso de mencionar “grande parte”). De realçar também o facto de grande parte dos grupos não interpretar as medidas de tendência central quando analisa os dados, ficando-se pela leitura de tabelas e gráficos.

Estes resultados são importantes porque mostram que há uma grande variedade de interpretações destas medidas, umas adequadas e outras não, maior ainda do que o indicado em investigações anteriores (e.g., Fernandes & Barros, 2005; Martins et al., 2009). Os futuros professores devem contactar e discutir todas elas, de modo a perceberem quais transmitem melhor a informação sobre os dados. É também

importante levá-los a comparar e contrastar as três medidas de tendência central de modo a perceber as respetivas diferenças. Por último, este estudo mostra a importância de realçar as interpretações concretas e contextualizadas nos dados para consolidar a compreensão das medidas, aprender a distingui-las e a perceber a sua utilidade.

Agradecimento

Este trabalho é financiado por fundos nacionais através da FCT – Fundação para a Ciência e Tecnologia no âmbito do Projeto *Desenvolver a literacia estatística: Aprendizagem do aluno e formação do professor* (contrato PTDC/CPE-CED/117933/2010).

Referências

- Batanero, C., & Godino, J. D. (2005). Perspectivas de la educación estadística como área de investigación. In R. Luengo (Ed.), *Líneas de investigación en Didáctica de las Matemáticas* (pp. 203-226). Badajoz: Universidad de Extremadura.
- Chatzivasileiou, E., Michalis, J., & Tsaliki, C. (2010). Elementary school students' understanding of concept of arithmetic mean. In C. Reading (Ed.), *Actas da 8th International Conference on Teaching Statistics*. Ljubljana, Eslovénia.
- Fernandes, J.A. (2009). Ensino e aprendizagem da Estatística: Realidades e desafios. *Actas do XIX EIEM: Números e Estatística*. Vila Real.
- Fernandes, J. A., & Barros, P. M. (2005). Dificuldades em estocástica de uma futura professora do 1º e 2º ciclos do ensino básico. *Revista Portuguesa de Educação*, 18(1), 117-150.
- Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, 70(1), 1-25.
- Groth, R. E. (2006). An exploration of students' statistical thinking. *Teaching Statistics*, 28(1), 17-21.
- Jacobbe, T. (2008). Elementary school teachers' understanding of the mean and median. In C. Batanero, G. Burrill, C. Reading & A. Rossman (Eds.), *Actas do ICMI Study 18 and 2008 IASE Round Table Conf*. Monterrey, México.
- Leavy, A., & O'Loughlin, N. (2006). Preservice teacher understanding of the mean: Moving beyond the arithmetic average. *Journal of Mathematics Teacher Education*, 9(1), 53-90.
- Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Mahwah, NJ: Erlbaum.
- Martins, C., Pires, M. V., & Barros, P. M. (2009). Conhecimento estatístico: Um estudo com futuros professores. *Actas do EIEM: Números e Estatística*. Vila Real.
- ME (1997). *Orientações curriculares para a educação pré-escolar*. Lisboa: DEB.
- ME (2007). *Programa de Matemática do ensino básico*. Lisboa: DGIDC.
- Mulekar, M. (2007). Preparing teachers of statistics in the United States. *Atas 56th Session of International Statistical Institute*. Lisboa.
- Pimentel, T., Vale, I., Freire, F., Alvarenga, D., & Fão, A. (2010). *Matemática nos primeiros anos: Tarefas e desafios para a sala de aula*. Lisboa: Texto.

- Ponte, J. P. (2007). Investigations and explorations in the mathematics classroom. *ZDM*, 39, 419-430.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Research*, 15(2), 4-14.
- Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67(3), 223-265.